

The Measure of Diffusion and Impact of Scientific Findings from the NHLBI-Supported Proteomics Centers Program (NPCP)

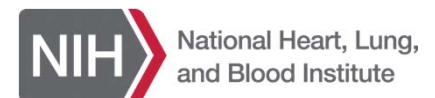
Tara S. Dasigi, MPH

NHLBI Summer Intern

National Heart, Lung, and Blood Institute - National Institutes of Health

Mentors: Dr. Richard Fabsitz, Dr. Pothur Srinivas, and Ms. Mona Puggal

August 5, 2014



Background & Introduction

- The NHLBI Proteomics Centers Program (NPCP) is a contract program comprised of seven university-based centers.
 - Last summer:
 - Analyzed the database based on three objectives:
 - Objective 1: To develop proteomic approaches and technologies,
 - Objective 2: To conduct molecular and mechanistic studies, and
 - Objective 3: To apply proteomic approaches and technologies to clinical questions.

Background & Introduction

- My project's goal:
 - To measure the impact and diffusion of current scientific findings (August 2010 – June 2013) from the NPCP.
 - Three major questions:
 - Has the Proteomics Program had an impact on the scientific community?
 - What factors predict impact?
 - Does impact vary by objective of the research?

Background and Introduction Cont'd

- Impact predictors in previous studies:
 - Number of authors: predicts high citation counts and impact of publications (Chen et al. & Kostoff et al.).
 - Journal importance (Chen et al.)
 - Author reputation: frequency of appearances of author-journal pairs during the four-year window prior to the investigation (Boyack et al.)
- *In the current project, multiple measures of impact and diffusion will be considered.*

Methods: Question 1 (Q1)

- Intent is to measure impact using five different approaches.
 - Percentile Ranks (PRs)
 - Scientific Uptake (# of citations)
 - Scientific Diffusion (SD)
 - Institutional Diffusion (ID)
 - Geographic Diffusion (GD)
- PRs: Focus for Q1 - understand if the NPCP has an impact on scientific community using impact and diffusion.
 - Defined: paper is in the top $n\%$ of papers in that category in that publication year (Thomson Reuters)

Methods: Question 2 (Q2)

- The Independent Variables/Impact Predictors:
 - Time since published (in months)
 - Publication type (PT) – e.g., original data, opinion, review, conference paper.
 - Collaborative disposition (CD) of the senior author
 - # of authors (AUTHORS)
 - # of institutions (INSTITUTIONS)
 - Journal Impact Factor (JIF)
 - Subject area (SAs).
- Outcome (Dependent Variable): # of cites (TIMESCIT).
- Regression Model:

$$\text{InTIMESCIT} = \text{AUTHORS} + \text{INSTITUTIONS} + \text{JIF} + \text{SA} + \text{Original data} + \text{Time} + \text{CD} + \text{ID}$$

Methods: Question 3 (Q3)

- In order to understand whether impact varies by objective:
 - Analyses done for Q2 repeated by adding the two indicator variables for mechanistic study and clinical study.
 - Created indicator variables for objective and coded as 0 or 1.
 - Reference category: Proteomic tools and technologies.

Results

- Biochemical research methods – most publications fall in this SA.

Table 3: Subject Areas and Their Respective Publications

| Subject Area (SA) | # of Publications |
|---|-------------------|
| Biochemical Research Methods | 99 |
| Biochemistry & Molecular Biology | 68 |
| Cardiac & Cardiovascular Systems | 61 |
| Clinical Neurology & Critical Care Medicine | 2 |

- Center A/Affiliated Colleges: most publications published by this institution

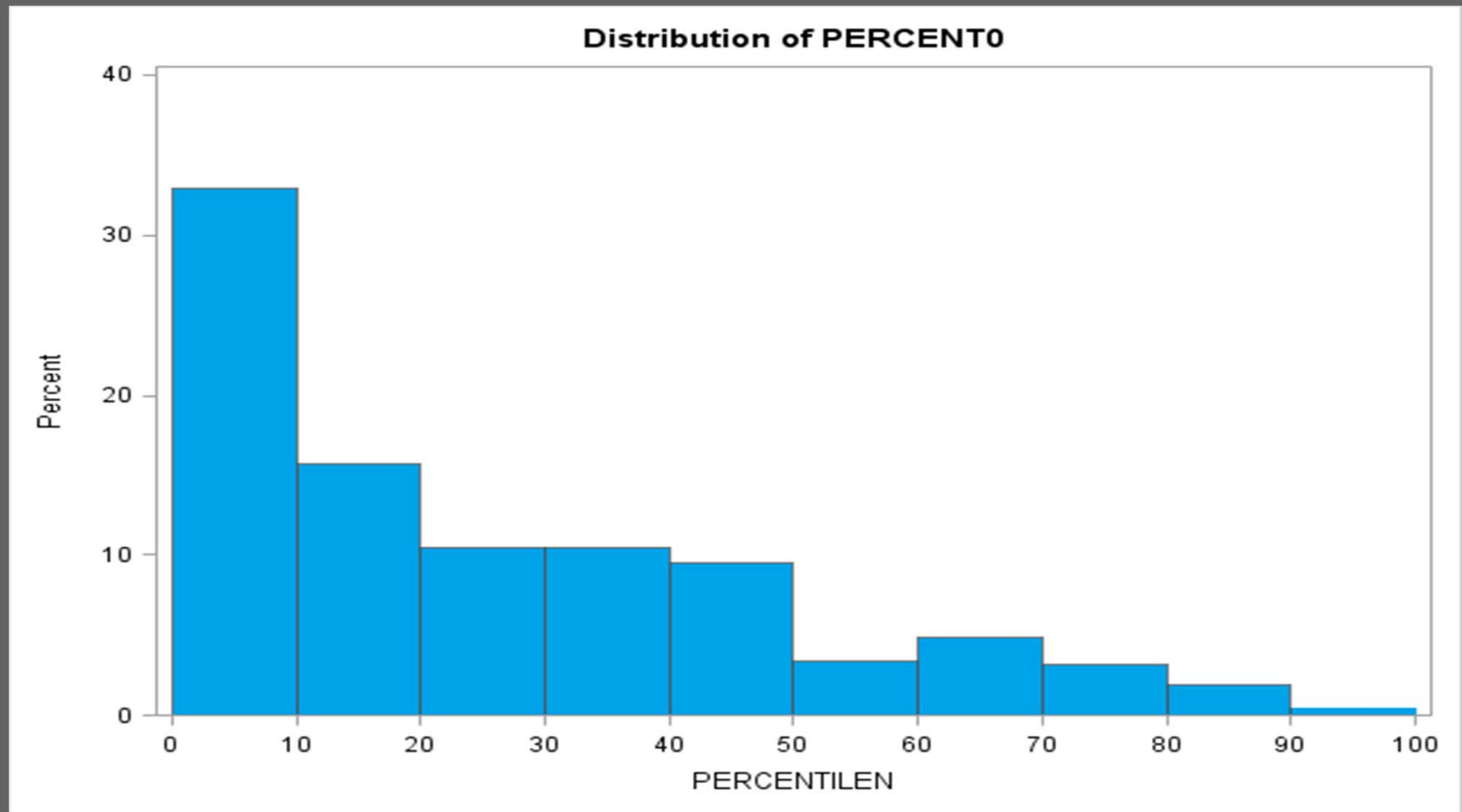
Table 4: Institutions and Their Respective Publications

| Center | # of Publications | Center | # of Publications |
|--------|-------------------|--------|-------------------|
| A | 112 | E | 42 |
| B | 76 | F | 33 |
| C | 75 | G | 25 |
| D | 65 | | |

*N = # of publications analyzed out of 463

Results

Figure 1: Percentile Rankings (PRs) Are Useful



*Lower PRs are better here

Predicting Impact - How Did We Do?

Table 3: Regression of Impact Predictors Excluding Objective

| Variable | Parameter Estimates | | |
|----------------|---------------------|----------------|------------------|
| | Parameter Estimate | Standard Error | Pr > t |
| Intercept | 0.28233 | 0.38336 | 0.4618 |
| AUTHORS | 0.01812 | 0.02023 | 0.3708 |
| INSTITUT | -0.01386 | 0.02853 | 0.6274 |
| <u>JIF</u> | <u>0.06476</u> | <u>0.01023</u> | <u><.0001</u> |
| Original | 0.33811 | 0.19711 | 0.0870 |
| DISPOSIT | -0.00368 | 0.02393 | 0.8779 |
| Institute | 0.09242 | 0.17822 | 0.6043 |
| Biochemical | 0.11838 | 0.36385 | 0.7451 |
| Biochemistry | -0.04092 | 0.37971 | 0.9142 |
| Engineering | 0.00283 | 0.45458 | 0.9950 |
| Cardiovascular | 0.02590 | 0.36699 | 0.9438 |
| Endocrinology | 0.09528 | 0.38985 | 0.8070 |
| Chemistry | -0.07463 | 0.44064 | 0.8656 |
| Immunology | -0.45826 | 0.39057 | 0.2413 |
| Genetics | -0.30247 | 0.47487 | 0.5245 |
| Microbiology | -0.44737 | 0.46230 | 0.3337 |
| MONTHS | 0.01427 | 0.00752 | 0.0585 |

ing,
te

What does a stepwise function do?

Table 4: Stepwise Regression of Impact Predictors Excluding Objective

| Variable | Parameter Estimate | Standard Error | Pr > F |
|---------------|--------------------|----------------|------------------|
| Intercept | 0.59226 | 0.20702 | 0.0044 |
| <u>MONTHS</u> | <u>0.01376</u> | <u>0.00697</u> | <u>0.0490</u> |
| <u>JIF</u> | <u>0.06485</u> | <u>0.00970</u> | <u><.0001</u> |

p -value: to enter = 0.05; to stay = 0.10; Force: Months

- Stepwise functions depict the reduced model required for predicting impact.

Predicting Impact – How Did We Do?

Table 5: Regression of Impact Predictors Including Objective

| Variable | Parameter Estimates | | |
|------------------|---------------------|----------------|------------------|
| | Parameter Estimate | Standard Error | Pr > t |
| Intercept | 0.62917 | 0.43739 | 0.1510 |
| AUTHORS | 0.01833 | 0.02014 | 0.3631 |
| INSTITUT | -0.00719 | 0.02858 | 0.8014 |
| <u>JIF</u> | <u>0.06462</u> | <u>0.01017</u> | <u><.0001</u> |
| Original | 0.35842 | 0.19604 | 0.0682 |
| DISPOSIT | -0.01372 | 0.02440 | 0.5743 |
| <u>Mechanism</u> | <u>-0.39176</u> | <u>0.23478</u> | <u>0.0959</u> |
| <u>Clinical</u> | <u>-4.15163</u> | <u>1.71254</u> | <u>0.0157</u> |
| Institute | 0.07176 | 0.17876 | 0.6883 |
| Biochemical | 0.07940 | 0.36225 | 0.8266 |
| Biochemistry | 0.00580 | 0.37793 | 0.9878 |
| Engineering | 0.02710 | 0.45174 | 0.9522 |
| Cardiovascular | 0.16488 | 0.36863 | 0.6549 |
| Endocrinology | 0.16463 | 0.38898 | 0.6723 |
| Chemistry | -0.12476 | 0.43937 | 0.7766 |
| Immunology | -0.41507 | 0.38873 | 0.2862 |
| Genetics | -0.40960 | 0.47639 | 0.3904 |
| Microbiology | -0.40320 | 0.45965 | 0.3809 |
| MONTHS | 0.01364 | 0.00748 | 0.0687 |

What does a stepwise function do?

Table 6: Stepwise Regression of Impact Predictors Including Objective

| Variable | Parameter Estimate | Standard Error | Pr > F |
|-----------|--------------------|----------------|--------|
| Intercept | 0.34047 | 0.24810 | 0.1706 |
| MONTHS | 0.01220 | 0.00695 | 0.0799 |
| JIF | 0.06522 | 0.00965 | <.0001 |
| Original | 0.37579 | 0.18944 | 0.0479 |
| Clinical | -3.48497 | 1.66144 | 0.0365 |

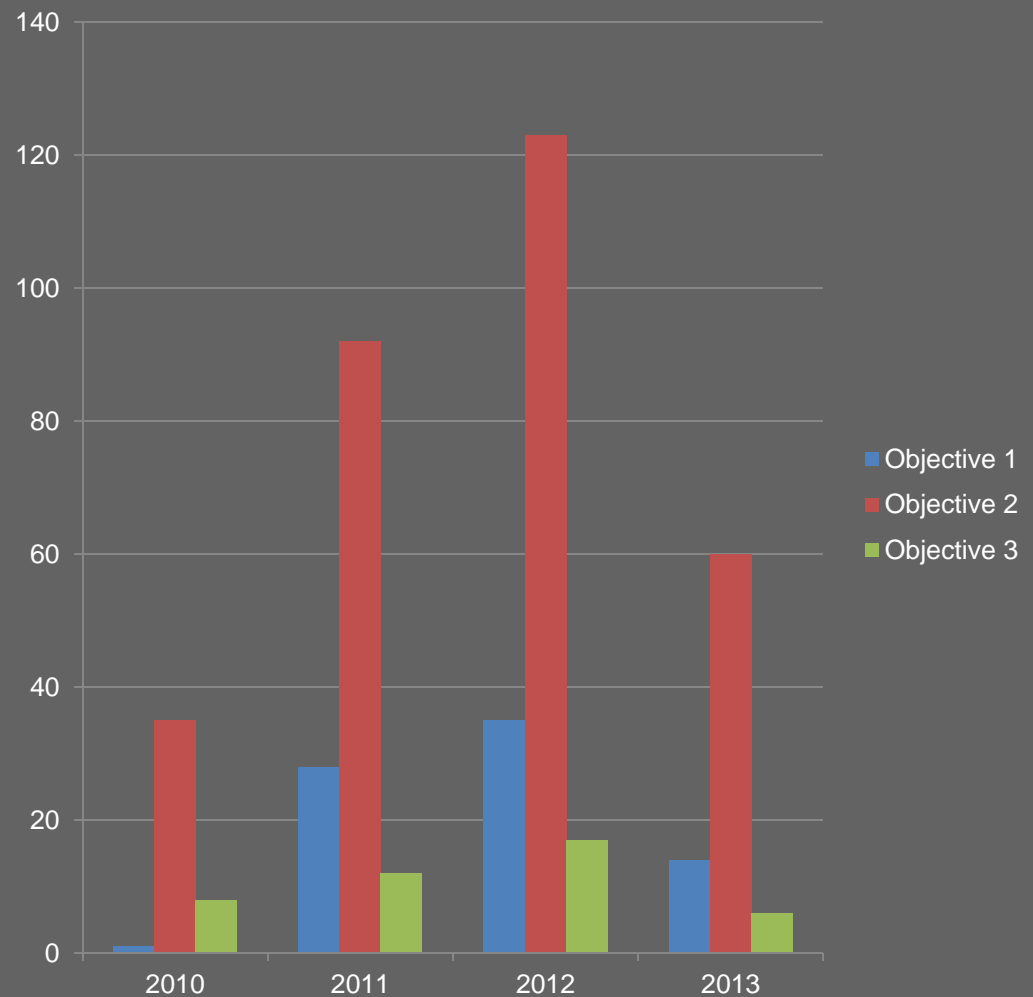
p -value: to enter = 0.05; to stay = 0.10; Force: Months

- Stepwise functions depict the reduced model required for predicting impact.

Results

- Citations for clinical studies were significantly lower compared to mechanistic studies and technological tool studies.

Figure 2: # of Publications by Objective per Year



Discussion and Conclusion

- NPCP has had a major impact on the scientific community.
 - PRs provide some insight on impact.
- Of all the putative predictors of impact (# of citations), JIF was the only significant predictor.
- When adding clinical and mechanistic studies, citations were significantly lower for clinical studies compared to their mechanistic and technological tool counterparts

In the future, clinical studies might overtake mechanistic studies with regards to having more citations.

Future Directions

- Diffusion: how many people have cited a publication of interest
 - Scientific Areas: Paper A published in Immunology is being cited by Paper B from Physics.
 - Breadth of SAs of the original publications and of those that have cited the publication.
 - Institutional (ID): Paper A has been cited by UCLA, Emory, and Cambridge.
 - Geographic (GD): Paper A in Baltimore, MD, has been cited in Los Angeles, CA; Athens, GA; and Toronto, ON.
 - $\text{Log}\sum(\text{all distances from institutions of the senior author of cited publication to institutions of the senior author of citing publication})$

Any questions?

Thank you!

Tara Dasigi, MPH

tdasigi@uga.edu



National Heart, Lung,
and Blood Institute